SELECTED PROBLEMS OF INTER-COUNTRY COMPARISONS ON THE BASIS OF THE EXPERIENCE OF THE EEC

BY DINO GERARDI

*Statistical Office of the European Communities**

The formidable expansion in the scope of the United Nations International Comparison Project has brought into evidence limitations of the methodology used in the first three phases. The author considers that there are two indispensable conditions needed to give renewed impetus to the ICP: (a) the objectives must be redefined, and (b) the methodology must be built on an entirely new basis. He considers the broad lines of such an evolution to be the following.

(a) The objective of volume comparison must be kept distinct from that of purchasing power comparison, given that both the basic material and the formulae to be used at the aggregate level differ in the two cases.

(b) At the basic heading level, it is proposed, for both volume and purchasing power comparisons, to replace the multilateral approach by a "minimum scale" binary and unilateral approach, and to use the EKS method. This will make possible an improvement in the accuracy of the estimates, a reduction in the overall costs, and a drastic reduction in execution time. What is more, it would be possible to regionalize the worldwide comparison, in the sense that the results of the basic heading comparisons already obtained at the regional level for regional purposes can be used as an input in the framework of the worldwide comparison. At the aggregate level, in the framework of volume comparison, it is proposed that a constant price procedure in the spatial sense should continue to be used. It is, however, proposed that the prices of the set of countries (GK) be replaced by a structure of common "equi-distant" prices (G). This would permit the elimination of the significant systematic distortions observed in the comparison between rich and poor countries in the first three phases of ICP. What is more, this gives maximum stability to results obtained for the same countries at different geographical levels. By using a set of common "equi-distant" quantities, the same advantage can be obtained in the purchasing power comparison.

INTRODUCTION

Owing to increasing demand for internationally comparable statistics, there has been a veritable boom in international comparisons of economic development levels (*per capita* gross domestic product, purchasing power) during the last two decades and particularly during the 1970s. These comparisons have been on a world scale, a regional scale (Europe, Latin America, etc.) and sub-regional scale (EEC, COMECON, etc.). As a comparison at a specific geographical level (worldwide, for example) includes comparisons that may already have been carried out at a lower level (regional or sub-regional, for example), it is understandable that the choices made at the various levels with regard to statistical concepts, methodology and organization should be interdependent. This unavoidable interdependence has given rise to a debate on the theory and practice of international comparisons, originally confined to workers and experts in the field and then extended to include other members of the scientific community. It is hoped that the outcome of this debate will be to produce basic guidelines

*The ideas expressed in this paper are the sole responsibility of the author and are not necessarily shared by the Statistical Office of the European Communities.

for a system of international comparisons at various geographic levels to be followed in the years to come.

This paper is my personal contribution to this debate.[1] The main ideas expressed in this paper, which are summarized below, are nearly always diametrically opposed to those which have inspired the United Nations International Comparison Project (ICP) up to the present time. Contrary to the beliefs underlying the ICP, the purchasing power parities (PPPs) used to carry out volume comparisons of GDP and its uses (and here a more suitable term might be "spatial deflators") cannot be considered to be genuine PPPs and hence to be instruments that enable us to compare, on the one hand, earnings in real terms and on the other, price levels in different countries, because both the raw material (prices and expenditures) and the formulae used are inadequate.

With regard to the choice of formula to calculate spatial deflators and volume ratios, my position is as follows: as an index reduces *de facto* the basket of commodities to a single theoretical composite article, and as every index attributes a specific degree of representativity to such a composite article, I am in favour of indices that conceive of the basket as a composite article that is "equi-representative" of countries (balanced indices). In my view both heterogeneous weighted indices that conceive of the basket as a typical composite article for the region made up of all the countries participating in the comparison, and homogeneous weighted indices—among which we may place the Geary–Khamis index used in the ICP—which conceive of the basket as a composite article whose degree of representativity is diametrically opposite, and which, by formal analogy, I have called an "atypical" article for the region, are therefore biased.

I prefer to use, among the balanced indices, the Elteto–Köves–Sculc (EKS) index at the basic headings classification level. Using this index leads naturally to a binary approach and hence to the regionalization of comparisons, as it generally means that the results calculated at a particular geographical level constitute the input for calculations at a higher geographical level. The binary approach, which is at an opposite pole from the multilateral approach used by ICP, guarantees that the estimates are more accurate and ensures that the comparisons would be less expensive to carry out. What is more, it encourages more countries to take part in the project, as it is much easier for a country to compare itself with a neighbouring country on the basis of clearly defined and representative commodities than to compare itself simultaneously with a whole group of countries from different areas in the world on the basis of broadly defined commodities. Statisticians know which commodities are recorded in their own country but they do not know which ones are recorded in other countries, and this is not just a matter for idle curiosity when a comparison is actually carried out.

At levels of aggregation higher than the basic headings, the "average test" of volume ratios requirement (i.e., the requirement that the volume ratio of an aggregate must be between the smallest and largest volume ratios of the com-

[1]The inspiration for my ideas has come mainly from the work of Drechsler, but also from that of Krzeczkowska, Köves and others. Some of the solutions I put forward were used by the EEC in its 1975 comparisons, while others are being used for its 1980 comparisons.

ponents) entails the abandonment of balanced indices and hence of the EKS index. The adoption of a "constant price" (in the spatial sense) procedure would however meet this requirement. In order for the basket of commodities (of basic headings, in this case) to be regarded as an equi-representative composite article, the system of common prices must be "equi-distant" from national price systems. If the composite article is not equi-representative, or if the common price system is not equi-distant, the results of the comparison will be biased.

There is another argument to illustrate what I am saying. Using common prices to measure the relative volumes of the various countries is rather like using an "elastic tape measure", the length of which is inversely proportional (given the negative correlation which normally exists between relative prices and relative quantities) to the "difference" between the system of common prices and the system of national prices (the Gershenkron effect). If world prices are used, for instance, the volumes of poor countries are inevitably overestimated because world prices are closer to those of rich countries. (A GK estimate of a volume ratio of a poor country relative to a rich country is almost always higher than an estimate of the Fisher type, which may be regarded as an "equi-distant constant price" procedure as we shall see.) The aim is therefore to achieve "equi-distance" between common prices and national prices in order to "stiffen" the tape measure. If we proceed in this way, differences in the results obtained at the various geographical levels for a given pair of countries will be minimal. What differences there are may be put down to random errors during the measuring process. (The result obtained at the world level may be closer to the binary result than, for example, the result achieved at the sub-regional level.)

Equi-distance in the system of common prices is in complete antithesis to the position adopted by ICP statisticians, who regard the use of prices for the region made up of all the countries (GK) as inevitable. The arguments put forward to defend this position are based on the conceptual error to the effect that the aims of a volume comparison and those of the aggregation of accounts of the various countries are in conflict. It is in fact easy to show that the value structure of the group of countries (that is, the result of the aggregation, given that absolute values depend on the arbitrary choice of the monetary unit used to express them) is given by the arithmetic mean of existing national structures weighted by overall volume ratios. (It should be noted that this applies irrespective of the method used to obtain them.)

Accordingly, we can say that, in order to obtain the best value structure for groups (not just the total) of countries, it is sufficient, but also necessary, to produce the best volume ratios. What is more, if regional prices are used, it is likely that there will be big differences in the results obtained at the various levels as—understandably—regional characteristics vary greatly from one geographical level to another. Therefore the idea that using the same index at various geographical levels promotes consistency of the results obtained at the various levels is correct if and only if the "common" index proposed is based on equi-representativity of the composite item.

For both comparisons of purchasing powers and volume comparisons I would propose the EKS index at the level of basic headings, as the requirements placed on the two types of comparison are the same at this level of aggregation. The "average test" of PPPs must be imposed on the results of the comparison

at levels of aggregation higher than the basic headings. I suggest that a "constant quantities" procedure should be used as an instrument to measure relative price levels in the various countries in order to meet this requirement. On the basis of the same arguments I used before, the objective must be to obtain common quantities that are "equi-distant" from national quantities.

1. Types of Comparison

1.1. *Volume Comparisons*

The principal objective of the work undertaken by the various international organizations (UN, EEC, CMEA, etc.) has been volume comparisons of gross product and its final uses (consumption, capital formation, government) among countries. The results of this type of comparison can be presented in any of the following forms:

- absolute volumes (or real values): i.e., nominal values converted by specific PPPs into a common reference unit;
- volume ratios: i.e., the absolute volumes of each country related to the absolute volumes of a reference country;
- relative parts: i.e., the absolute volumes of each country related to the absolute volumes of all participating countries, obtained by addition.

With regard to the uses of these results, real *per capita* gross domestic product (GDP) and its principal components are among the most important ingredients of several kinds of economic analysis and projection. In certain regions (EEC, CMEA), real GDP per head of the different countries and its evolution over time allow one, on the one hand, to establish the conditions, and on the other hand, to follow the development of the economic integration of the countries. The geographic distribution of real GDP and other important aggregates within a region (the contribution of each country to the regional aggregates) also enables one to derive weighting schemes for the aggregation of national economic indicators (e.g. temporal deflators) in order to obtain these indicators for the region as a whole.

The geographic distribution of the real GDP of a region is also used to distribute regional GDP into its final use components (consumption, capital formation, government). Such a structure is an average of the national structures, weighted by the relative shares of the national real GDPs in the GDP of the region (whatever method is used). This means that in order to obtain the best value structure for groups and sub-groups of countries it is not only sufficient but also absolutely necessary to produce the best possible volume comparisons. The relative shares of the national real GDPs in the GDP of the region can also be used to establish in an equitable way the contributions of each country to the budget of supra-national organizations.

The volume ratios can be calculated directly by estimating the quantities consumed or invested in the different countries, or indirectly by eliminating the component prices from the ratios of the nominal values. Generally, the latter method is used, for reasons not discussed here of theoretical and practical opportunity. The basic data for the indirect calculation of volume ratios are therefore national nominal values. An essential condition is that these nominal

values be comparable, i.e., derived from harmonized systems of national accounts. To complete the data base, average annual national prices for a sample of items in the different aggregates must be estimated. These prices must be of the same type as those embodied in the nominal values. They must, for example, include subsidies if these are included in the nominal values. Dividing the ratios of the nominal values by the PPPs would then yield volume ratios that constitute the final result of the comparison.

1.2. *Purchasing Power Comparisons*

The comparison of the purchasing power of currencies has to date been considered a secondary objective by the main international organizations (UN, EEC, etc.). For some specific areas the PPPs obtained as a by-product of volume comparisons may provide useful information when treated with sufficient caution, but in general the results are completely unusable. They may be both highly misleading due to inadequacy of the basic data used, and unacceptable because they fail to meet some important formal condition, as for instance the average test.[2]

The main results from this type of comparison are PPPs between the currencies of the participating countries. These express the ratios between the average levels of prices for items in a given group (investment goods, foodstuffs). By relating the PPPs for specific items to the PPPs for larger groups of items, comparisons of the relative price levels can be made between countries. Usually, PPPs for a given group of items are then related to exchange rates, and conclusions are drawn on the reciprocal over- or under-valuation of the various currencies. Such conclusions must, however, be drawn with a certain amount of caution. PPPs are also widely used to make international comparisons of the purchasing power of salaries.

Prices used in this context are those effectively paid by the purchaser; they therefore should exclude subsidies. Weighting factors are derived from market flows, i.e., quantities consumed or invested multiplied by purchase prices. Often the weighting factors are distributions in percentage terms. In order to obtain these data, only minor modification of the data already available for volume comparison is needed. Given the importance of purchasing power comparisons and the relatively low financial cost involved, the best solution for the future would be to raise this type of comparison to the level of an independent study.

2. Methods

2.1. *Index Number Formulae*

The indices proposed in the specialist literature are so varied that any attempt to list them or to carry out an exhaustive and meaningful classification

[2]For example in Phase II of ICP the following PPPs against the U.S. dollar were obtained for 1973:

Aggregate	*Netherlands*	*France*
Government compensation	2.87	3.62
Government commodities	3.18	1.117
Government total	3.23	4.21

would certainly be doomed to failure. We have therefore elected to present six groups of indices which, in addition to incorporating many of the better known indices, also provide in our view a sufficiently wide range of possibilities in current applications. As each of these groups contains a Van Yzeren index, we have called each group by the name of the corresponding Van Yzeren index. This gives us a group of homogeneous non-weighted (O) and weighted (OW) indices, one of heterogeneous non-weighted (E) and weighted (EW) indices, and one of balanced non-weighted (B) and weighted (BW) indices.[3] Each of these groups is, moreover, made up of G-type and VY-type indices; the G-type indices imply a common basket of items or basic headings.

2.1.1. Homogeneous Non-Weighted Indices

(*a*) *Indices of the* $(X)VY(O)$ *type.* The purchasing power parity (PPP) between the monetary unit of a hypothetical reference country and that of country h can be expressed, using indices of this type, as follows:

$$(1) \qquad {}^{(X)}_{h}P^{VY(O)} = \left[\sum_{\alpha=1}^{k} ({}^{(X)}_{\alpha}P^{VY(O)}\,{}_{h}P^{L}_{\alpha})^{X}/k\right]^{1/X} \qquad h = 1, 2, \ldots, k$$

in which ${}_{h}P^{L}_{\alpha}$ is the PPP of the Laspeyres type[4] between country α and country h, and k is the number of countries. The PPPs between the currencies of the various countries concerned can therefore be deduced from the previous ones as follows:

$$(2) \qquad {}^{(X)}_{h}P^{VY(O)}_{j} = {}^{(X)}_{h}P^{VY(O)}/{}^{(X)}_{j}P^{VY(O)} \qquad \forall(j, h).$$

The only indices of the $(X)VY(O)$ type that are of interest by virtue of their interpretation and the properties they satisfy are those obtained by $X = 1$, $X = -1$, and calculating $\lim X \to 0$. The first two are made up of systems of equations that are easily solved iteratively and are characterized by interpretations of the "tourist"[5] and "migratory"[6] types respectively, while the third consists of the following explicit form:

$$(3) \qquad {}^{(0)}_{h}P^{VY(O)}_{j} = \left[\prod_{\alpha=1}^{k} {}_{h}P^{L}_{\alpha}/{}_{j}P^{L}_{\alpha}\right]^{1/k} \qquad \forall(j, h)$$

and is characterized, as can be seen from (3), by an interpretation of the "bridge country" type.[7] The three indices coincide with the Fisher index when only two countries are considered.

[3]The weighting factors referred to here, which distinguish the homogeneous, heterogeneous and balanced weighted indices from the corresponding non-weighted indices, are proportional to the size of the countries (country weights), and must not be confused with the weightings that are proportional to the significance of the individual items (item weights) which, except in certain rare cases, are a feature of all the indices used in the various applications.

[4]It is assumed that the reader is familiar with the Laspeyres type of price and volume index, and also with the Paasche and Fisher indices.

[5]Van Yzeren, "Three Methods of Comparing the Purchasing Power of Currencies," Netherlands Central Bureau of Statistics, Statistical Studies (December 1956), pp. 3–34.

[6]Gerardi, D., "Sul problema della comparazione dei poteri d'acquisto delle valute," Istituto di Statistica dell'Universita di Padova, Serie papers (August 1974).

[7]C. Gini, "On the Circular Test of Index Numbers," *Metron* (February 1931).

(*b*) *Indices of the* $(X)G(O)$ *type.*[8] The ${}^{(X)}_{h}P^{G(O)}$ can be expressed as follows:

$$(4) \qquad {}^{(X)}_{h}P^{G(O)} = \sum_{i=1}^{N} \pi_i q_{ih} \Big/ \sum_{i=1}^{N} p_{ih} q_{ih}$$

$$\pi_i = \left[\sum_{\alpha=1}^{k} ({}^{(X)}_{\alpha}P^{G(O)} p_{i\alpha})/k \right]^{1/X}$$

in which $p_{i\alpha}$ and $q_{i\alpha}$ represent the price and quantity respectively of item i in country α, N the total number of items or basic headings of a common classification, and π_i the prices in the hypothetical reference country. The indices obtained from (4) by taking $X = 1$ and calculating $\lim X \to 0$ are the only ones of any relevance. The first comprises a system of equations that is easily solved iteratively and coincides, in the case of a common list of items (basic headings), with the $(1)VY(O)$ index. Thus when only two countries are considered the $(1)G(O)$ index provides a new and interesting expression of the Fisher index. The second is an explicit form and is the one which, as we shall see, will prove the most useful from the point of view of properties satisfied.[9] It is written as follows:

$$(5) \qquad {}^{(0)}_{h}P^{G(O)} = \sum_{i=1}^{N} \left[\prod_{\alpha=1}^{k} p_{i\alpha} \right]^{1/k} q_{ih} \Big/ \sum_{i=1}^{N} p_{ih} q_{ih}$$

This index, which was used for the first time by the EEC for the 1975 volume comparison and tested in the third phase of the ICP for world-level volume comparison, is the one we propose as the basic method of aggregation in volume comparisons within the framework of the ICP, for reasons that will be explained later.

2.1.2. Homogeneous Weighted Indices

(*a*) *Indices of the* $(X)VY(OW)$ *type.* These indices are obtained simply by incorporating country weights (W_α) in expression (1):

$$(6) \qquad {}^{(X)}_{h}P^{VY(OW)} = \left[\sum_{\alpha=1}^{k} ({}^{(X)}_{\alpha}P^{VY(OW)}{}_{h}P^{L}_{\alpha})^{X} w_\alpha \Big/ \sum_{\alpha=1}^{k} w_\alpha \right]^{1/X}$$

$$w_\alpha = {}^{(X)}_{\alpha}P^{VY(OW)} \sum_{i=1}^{N} p_{i\alpha} q_{i\alpha}.$$

The only indices of the $(X)VY(OW)$ type that are of interest are those obtained by $X = 1$, $X = -1$, and calculating $\lim X \to 0$. The first and third are made up of systems of equations that are easily solved iteratively and are characterized by interpretations of the "tourist" and "bridge country" types respectively. The latter takes the following explicit form:

$$(7) \qquad {}^{(-1)}P^{VY(OW)}_{h} = \sum_{\alpha=1}^{k} {}_{\alpha}P^{P}_{h} \sum_{i=1}^{N} p_{i\alpha} q_{i\alpha}$$

[8]Eurostat, "Comparison in Real Values of the Aggregates of ESA 1975," 1977.

[9]Gerardi, D., "Alcuni aspetti metodologici reguardanti il calcolo delle PPA fra i paesi della Comunita Europea," *Atti del Convegno della Societa Italiana di Statistica*, March 1978.

and is characterized by an interpretation of the migratory type. By using indirect parities of the ${}_h\bar{P}_j^{O(\alpha)} = {}_hP_\alpha^L/{}_jP_\alpha^L = {}_\alpha P_j^P/{}_\alpha P_h^P$ type and country weights of the $w'_\alpha = w_\alpha / \sum_{r=1}^{k} w_r$ type the $(O)P(OW)$ can be written as follows:

$$ {}_h^{(0)}P_j^{VY(OW)} = \prod_{\alpha=1}^{k} [{}_h\bar{P}_j^{O(\alpha)}]^{w'_\alpha}. $$

It can be seen that if country weights not dependent upon unknown PPPs (e.g., populations) were used, this index would be the only explicit form.

(*b*) *Indices of the* $(X)G(OW)$ *type.* This type of index is also obtained by incorporating weighting factors (W_α) into expression (4) above:

$$ (8) \qquad {}_h^{(X)}P^{G(OW)} = \sum_{i=1}^{N} \left[\sum_{\alpha=1}^{k} ({}_\alpha^{(X)}P^{G(OW)} p_{i\alpha})^X w_\alpha \Big/ \sum_{\alpha=1}^{k} w_\alpha \right]^{1/X} q_{i\alpha} \Big/ \sum_{i=1}^{N} p_{ih} q_{ih} $$

$$ w_\alpha = {}_\alpha^{(X)}P^{G(OW)} \sum_{i=1}^{N} p_{i\alpha} q_{i\alpha}. $$

Thus, in the case of the common list of items (basic headings), it becomes:

$$ {}_h^{(1)}P^{VY(OW)} = {}_h^{(1)}P^{G(OW)}. $$

Using country weights not dependent upon unknown PPPs, index $(0)G(OW)$, the weighted version of the index used by the EEC, becomes an explicit form. Indices $(1)G(OW)$ and $(0)G(OW)$ are the only indices of the $G(OW)$ type that are of any interest.

(*c*) *The Geary–Khamis index.* In the version originally proposed by Geary, the ${}_hP^{GK}$ can be written as follows:

$$ (9) \qquad {}_hP^{GK} = \sum_{i=1}^{N} \pi_i^{GK} q_{ih} \Big/ \sum_{i=1}^{N} p_{ih} q_{ih} \qquad h = 1, 2, \ldots, k $$

$$ \pi_i^{GK} = \sum_{\alpha=1}^{k} {}_\alpha P^{GK} p_{i\alpha} q_{i\alpha} \Big/ \sum_{\alpha=1}^{k} q_{i\alpha} \qquad i = 1, 2, \ldots, N. $$

The specific quantities $q_{i\alpha}$ in the definition of π_i^{GK} serve as country weights just like the W_α factors in (8), while the q_{ih} in the definition of ${}_hP^{GK}$ serve as item weights. The *GK* index was used as the basic method for aggregations in volume comparisons in the first three phases of the ICP. The results of the international comparison undertaken during phase III of the ICP using the *GK* index were compared with those obtained applying to the same basic data the $(0)G(OW)$ index, i.e., the weighted version of the index used by the EEC. The virtually identical results arrived at matched expectations.

2.1.3. Heterogeneous Non-weighted Indices

(*a*) *Indices of the* $(X)VY(E)$ *type.* The PPP between the monetary unit of country h and that of a hypothetical reference country can be expressed as follows:

$$ (10) \qquad {}^{(X)}P_h^{VY(E)} = \left[\sum_{\alpha=1}^{k} ({}^{(X)}P_\alpha^{VY(E)} {}_\alpha P_h^L)^X / k \right]^{1/X}. $$

The only indices of the $(X)VY(E)$ type that are of interest from the point of view of their interpretation and the properties they satisfy are those obtained using $X=1$, $X=-1$, and calculating $\lim X \to 0$. The first two are composed of systems of equations that are easily solved iteratively and are characterized by an interpretation of the "tourist" and "migratory" types respectively, while the third takes the following explicit form:

$$(11) \qquad {}^{(0)}_{h}P_{j}^{VY(E)} = \left[\prod_{\alpha=1}^{k} {}_{h}P_{\alpha}^{P}/{}_{j}P_{\alpha}^{P}\right]^{1/k}$$

and is characterized by an interpretation of the "bridge country" type. As can be seen, the three indices coincide with the Fisher index when only two countries are considered.

(*b*) *Index of the* $(X)G(E)$ *type.* The ${}^{(X)}P_{h}^{G(E)}$ can be expressed as follows:

$$(12) \qquad {}^{(X)}P_{h}^{G(E)} = \sum_{i=1}^{N} p_{ih}\left[\sum_{\alpha=1}^{k} \left(q_{i\alpha}/{}^{(X)}_{\alpha}P^{G(E)} \sum_{i=1}^{N} p_{i\alpha}q_{i\alpha}\right)^{X}/K\right]^{1/X}.$$

The factors

$${}^{(X)}q_{i}^{G(E)} = \left[\sum_{\alpha=1}^{K} \left(q_{i\alpha}/{}^{(X)}_{\alpha}P^{G(E)} \sum_{i=1}^{N} p_{i\alpha}q_{i\alpha}\right)^{X}/K\right]^{1/X}$$

represent the quantities of the hypothetical reference country. They are averages of the national quantities, standardized beforehand by means of the overall "volume". In the case of a common list of items (basic headings), this gives:

$${}^{(1)}P_{h}^{VY(E)} = {}^{(1)}P_{h}^{G(E)}.$$

Thus when only two countries are considered the $(1)G(E)$ constitutes a new and interesting expression of the Fisher index. Index $(0)G(E)$ is the only explicit form, and is the one that is the most useful in terms of the conditions satisfied. It is written as follows:

$$(13) \qquad {}^{(0)}P_{h}^{(E)} = \sum_{i=1}^{N} p_{ih}\left[\prod_{\alpha=1}^{K} q_{i\alpha}\right]^{1/k}.$$

This index is the one we advocate, for reasons that will be outlined later, for aggregations in comparisons of purchasing power within the framework of ICP.

2.1.4. Heterogeneous Weighted Indices

(*a*) *Indices of the* $(X)VY(EW)$ *type.* Introducing the weighting factors W_{α} in (10) gives:

$$(14) \qquad {}^{(X)}P_{h}^{VY(EW)} = \left[\sum_{\alpha=1}^{k} ({}^{(X)}P_{\alpha}^{VY(EW)}{}_{\alpha}P_{h}^{L})^{X}W_{\alpha}\Big/\sum_{\alpha=1}^{k} W_{\alpha}\right]^{1/X}$$

$$W_{\alpha} = {}^{(X)}_{\alpha}P^{VY(EW)}) \sum_{i=1}^{N} p_{i\alpha}q_{i\alpha}.$$

The only indices of the $(X)VY(EW)$ type that are of interest are those obtained by $X=1$, $X=-1$ and calculating $\lim X \to 0$. Index $(1)VY(EW)$ is the only

explicit form:

$$(15) \qquad {}^{(1)}P_h^{VY(EW)} = \sum_{\alpha=1}^{k} {}_{\alpha}P_h^L \sum_{i=1}^{N} p_{i\alpha}q_{i\alpha}.$$

Using the indirect parities of the ${}_h\bar{P}_j^{E(\alpha)} = {}_{\alpha}P_j^L/{}_{\alpha}P_h^L = {}_hP_{\alpha}^P/{}_jP_{\alpha}^P$ type, and country weights of the $W'_{\alpha} = W_{\alpha}/\sum_{r=1}^{k} W_r$ type, the ${}^{(0)}_hP_j^{(EW)}$ can be written as follows:

$${}^{(0)}_hP_j^{VY(EW)} = \prod_{\alpha=1}^{k} [{}_h\bar{P}_j^{E(\alpha)}]^{W'_{\alpha}}.$$

If country weights not dependent on unknown PPPs (e.g., populations) were to be used, then this index would be the only explicit form.

(*b*) *Indices of the* $(X)G(EW)$ *type.* Incorporating the weighting factors W_{α} into expression (12) gives:

$$(16) \qquad {}^{(X)}P_h^{G(EW)} = \sum_{i=1}^{N} p_{ih}\left[\sum_{\alpha=1}^{k}\left(q_{i\alpha}/{}^{(X)}_{\alpha}P^{G(EW)}\sum_{i=1}^{N} p_{i\alpha}q_{i\alpha}\right)^X W_{\alpha}\Big/\sum_{\alpha=1}^{k} W_{\alpha}\right]^{1/X}$$

$$W_{\alpha} = {}^{(X)}_{\alpha}P^{G(EW)}\sum_{i=1}^{N} p_{i\alpha}q_{i\alpha}.$$

Index $(1)G(EW)$, which in the case of a common list of items (basic headings) coincides with $(1)VY(EW)$, is the only explicit form.[10]

$$(17) \qquad {}^{(1)}P_h^{G(EW)} = \sum_{i=1}^{N} p_{ih}\sum_{\alpha=1}^{k} q_{i\alpha}.$$

This is the index used by the UN Economic Commission for Latin America.[11] Using country weights not dependent on the unknown PPPs, index $(0)G(EW)$ becomes an explicit form. Indices $(1)G(EW)$ and $(0)G(EW)$ are the only indices of the $G(EW)$ type that are of interest.

2.1.5. Balanced Non-Weighted Indices

(*a*) *Indices of the* $(X)VY(B)$ *type.* The ${}^{(X)}P^{VY(B)}$ can be expressed as follows:

$$(18) \quad {}^{(X)}_hP^{VY(B)} = \left\{\left[\sum_{\alpha=1}^{k}({}^{(X)}_{\alpha}P^{VY(B)}{}_hP_{\alpha}^L)^X/k\right]^{1/X}\Big/\left[\sum_{\alpha=1}^{k}({}^{(X)}P_{\alpha}^{VY(B)}{}_{\alpha}P_h^L)^X/k\right]^{1/X}\right\}^{1/2}.$$

Although in the formal expression the $(X)VY(B)$ indices are a simple geometric average of the $(X)VY(O)$ and $(X)VY(E)$ indices, the results they give rise to are not the same as the geometric average of the results obtained by applying the $(X)VY(O)$ and $(X)VY(E)$ indices separately, except for $(0)VY(B)$. The indices obtained from (18) by $X = 1$, $X = -1$ and calculating $\lim X \to 0$ are the only ones that are of interest, and they are characterized by an interpretation

[10] Gini, C., op. cit.

[11] U.N. Economic Commission for Latin America, "A Measurement of Price Levels and the Purchasing Power of Currencies in Latin America, 1960–1962," Document *E/CN.*12/653 (New York: United Nations, 1963).

of the "tourist," "migratory," and "bridge country," types respectively. The PPPs derived from index $(0)VY(B)$, known in the literature as the *EKS* index (from the initials of Elteto, Köves and Szulc), can be expressed as follows:

$$(19) \qquad {}^{(0)}_{h}P^{VY(B)} = [{}^{(0)}_{h}P^{VY(O)} \cdot {}^{(0)}_{h}P^{VY(E)}]^{1/2}.$$

The PPPs between countries j and h can be written in full as follows:

$$(20) \quad {}^{(0)}_{h}P^{VY(B)}_{j} = \left\{\left[\prod_{\alpha=1}^{k} {}_{h}P^{L}_{\alpha}\right]^{1/k}\left[\prod_{\alpha=1}^{k} {}_{h}P^{P}_{\alpha}\right]^{1/k}\right\}^{1/2} \Big/ \left\{\left[\prod_{\alpha=1}^{k} {}_{j}P^{L}_{\alpha}\right]^{1/k}\left[\prod_{\alpha=1}^{k} {}_{j}P^{P}_{\alpha}\right]^{1/k}\right\}^{1/2}$$

$$= \left[\prod_{\alpha=1}^{k} ({}_{h}P^{L}_{\alpha} \cdot {}_{h}P^{P}_{\alpha})^{1/2}\right]^{1/k} \Big/ \left[\prod_{\alpha=1}^{k} ({}_{j}P^{L}_{\alpha} \cdot {}_{j}P^{P}_{\alpha})^{1/2}\right]^{1/k}$$

$$= \left[\prod_{\alpha=1}^{k} {}_{h}P^{F}_{\alpha}/{}_{j}P^{F}_{\alpha}\right]^{1/k} = \left[\prod_{\alpha=1}^{k} {}_{h}\bar{P}^{B(\alpha)}_{j}\right]^{1/k} = \left[\prod_{\alpha=1}^{k} ({}_{h}\bar{P}^{O(\alpha)}_{j} \cdot {}_{h}\bar{P}^{E(\alpha)}_{j})^{1/2}\right]^{1/k}.$$

Expression (20) shows that the PPP between countries j and h is obtained as the simple geometric average of "indirect" PPPs of the ${}_{h}\bar{P}^{B(\alpha)}_{j}$ type, each of which uses country α as a bridge (each of the K countries being considered in turn as a bridge country). It is known that the PPPs obtained using the EKS index differ minimally from PPPs of the Fisher type:

$$\Delta = \sum_{r}\sum_{s} (\log {}_{s}EKS_{r} - \log {}_{s}F_{r})^{2} = \min.$$

The indices of the $(X)VY(B)$ type coincide with the Fisher index when only two countries are considered. The *EKS* index is the one we advocate, for reasons explained later, as the basic method for calculating the PPPs for the basic headings in both volume and purchasing power comparisons in the context of the ICP.

(*b*) *Indices of the* $(X)G(B)$ *type.* The ${}^{(X)}_{h}P^{G(B)}$ can be expressed as follows:

$$(21) \quad {}^{(X)}_{h}P^{G(B)} = \left\{\left(\sum_{i=1}^{N} p_{ih}q_{ih}\right)^{-1} \sum_{i=1}^{N}\left[\sum_{\alpha=1}^{k} ({}^{(X)}_{\alpha}P^{G(B)}p_{i\alpha})^{X}/k\right]^{1/x} q_{ih} \Big/ \right.$$

$$\left. \sum_{i=1}^{N} p_{ih}\left[\sum_{\alpha=1}^{k}\left(q_{i\alpha}/{}^{(X)}_{\alpha}P^{G(B)} \sum_{i=1}^{N} p_{i\alpha}q_{i\alpha}\right)^{X}/k\right]^{1/x}\right\}^{1/2}$$

In the case of a common list of items (basic headings) we have:

$${}^{(1)}_{h}P^{VY(B)} = {}^{(1)}_{h}P^{G(B)}$$

The following is also valid for $(0)G(B)$: ${}^{(0)}_{h}P^{G(B)} = [{}^{(0)}_{h}P^{G(O)} \cdot {}^{(0)}_{h}P^{G(E)}]^{1/2}$, $(0)G(B)$ being the only index of the $(X)G(B)$ type that takes an explicit form.

2.1.6. Balanced Weighted Indices

Combining $(X)VY(OW)$ indices with $(X)VY(EW)$ indices, and the $(X)G(OW)$ indices with the $(X)G(EW)$ indices, would give indices $(X)VY(BW)$ and $(X)G(BW)$ respectively. Although indices of these types have

been proposed by some authors[12] and although others[13] hint at the possibility of applying them in practice, it is no exaggeration to state that they amount to nonsense. This will emerge clearly in the next section when we come to analyse the degree of representativity of the indices examined above. In fact we can see that the representativity of the balanced weighted indices and that of the corresponding balanced unweighted indices are the same, which implies that the results obtained with the two kinds of indices would be almost identical.

2.2. *Representativity of the Composite Item*

The arguments put forward in this section are based on the twofold observation (1) that an index of prices or quantities *de facto* reduces a basket of individual items to a single (notional) composite item, and (2) that the adoption of an index is tantamount to giving that composite item a given type of representativity. An attempt is therefore made in this section to establish what type of representativity is conferred by the most familiar indices on the basket of items considered as one notional composite item. The choice of an index is thus brought down to the same level as the choice of an article. This simplification, despite its inherent limitations, makes it possible to establish a link between the abstract world of formulae and the concrete world of practical decision making, which is in my view an essential condition for making sound choices in the various stages of comparison: choice of items in the basket, choice of the index (price and volume) for the basic headings, and choice of the index (price and volume) for the aggregates.

2.2.1. Preliminary Remarks

It is familiar to everyone that, for a given pair of countries, the price ratio based on an item the consumption of which is higher in the denominator country will tend to be higher than that based on an item the consumption of which is higher in the partner country. To put it in a more learned manner, relative prices have a negative correlation with relative quantities. Furthermore, it is known that the use of the quantities consumed in the denominator country (the Laspeyres price index) will tend to result in a higher value than the Paasche index, which uses the quantities consumed in the partner country. On the basis of these considerations, it can therefore be said that the Laspeyres and Paasche price indices reduce the basket to a single composite item that is comparatively more representative, respectively, of the denominator country and the partner country. Since the Fisher index is a non-weighted (equi-weighted) geometric average of the Laspeyres and Paasche indices, it can be said to reduce the basket to a composite item that is equi-representative of both countries. This is confirmed by the fact that the quantities used by this index, considered as a binary version of the (1)$G(E)$ index, consist of a simple arithmetic average of the quantities of the two countries, standardized beforehand (using the respective overall volumes).

[12]Van Yzeren, *loc. cit.*

[13]I. B. Kravis, Kenessey, Zoltan, Heston, Alan W., and Summers, Robert, *A System of International Comparisons of Gross Product and Purchasing Power*, Johns Hopkins University Press, 1975.

2.2.2. The Typical Item

Let us now consider index $(1)G(EW)$. The parity ${}^{(1)}{}_{h}P_{j}^{G(EW)}$ between any pair of countries j and h is calculated by relating the quantities of the region comprising the countries concerned, multiplied by the prices of country j, to the same quantities multiplied by the prices of h. It can therefore be said that this index reduces the basket to a typical composite item for that region. In terms of relative representativity between countries, this means that the more the consumption pattern of a country resembles that of the region, the more the composite item will be representative of that country and vice versa. In the extreme case in which the quantities, e.g. of denominator country h, are proportional to those of the region, the ${}^{(1)}{}_{h}P_{j}^{G(EW)}$ coincides with the Laspeyres index, so that the composite item provides for that pair of countries and for that specific basket of individual items the highest degree of representativity in the base country. The same can be said for the indices $(0)VY(EW)$ and $(1)VY(EW)$, which coincide with $(1)G(EW)$ in the case of a common basket of items, in view of the substantial similarity of their make-up, and of the results which these indices normally provide in practice. This concept of the basket of items as a composite item typical of the region conforms to the very widespread tendency in the context of volume comparisons to select the items most consumed in the region to obtain a common basket of individual items. When only two countries are considered, this entails selecting the items most consumed in the bigger country.

2.2.3. The Atypical Item

Let us consider the GK index. The volume index ${}_{h}Q_{j}^{GK}$ between countries j and h is calculated by relating the quantities of j, multiplied by the prices of the region comprising the countries together, to the quantities of h multiplied by the same prices. The parity is expressed as the ratio between the nominal value index and the volume index:

$$ {}_{h}P_{j}^{GK} = {}_{h}V_{j} / {}_{h}Q_{j}^{GK} $$

It is clear that the closer the price system of one of the two countries is to that of the region, the relatively more representative of the partner country the composite item on which the index is based will be. If the prices of $j(h)$ tend to be proportional to those of the region, ${}_{h}Q_{j}^{GK}$ will tend towards index ${}_{h}Q_{j}^{P}({}_{h}Q_{j}^{L})$, and ${}_{h}P_{j}^{GK}$ will tend towards index ${}_{h}P_{j}^{L}({}_{h}P_{j}^{P})$, so that the composite item will tend to be relatively more representative of country $h(j)$. If the price system of one of the two countries is relatively closer to that of the region, the same will very likely be true of its consumption pattern as well. Insofar as this holds true, it can therefore be maintained that the GK index sees the basket as a composite item that will be relatively more representative of a given country, the closer the consumption pattern of the partner country is to that of the region. This is very far from the regional typicality described earlier. We can therefore say in the interests of formal symmetry that the GK index sees the basket as an atypical composite item for the region. This type of representativity given to the basket by the GK index is not reflected, in the general case of more than two countries,

in any strategy as regards selection of individual items for the basket. When only two countries are considered, the strategy of item selection consistent with the *GK* index (selection of atypical items for the region) consists of choosing the item most widely encountered in the . . . smaller country! This will certainly come as an unpleasant surprise for many of those who advocate the *GK* index, particularly in the context of the ICP. In fact, the item selection strategy largely favoured in this context, and also due to the use (or abuse) of the principle of "transactions equality," is precisely the opposite one of regional typicality which, in the case of two countries, only results in the selection of the items most widely encountered in the bigger country. The remarks made concerning the *GK* index also hold true to a very large extent for indices (1)$G(OW)$ and (0)$G(OW)$, because, on the one hand, of the great structural resemblance among the three indices, and on the other hand, of the high degree of equivalence in the results obtained using these indices in experiments.

2.2.4. The Equi-Representative Item

In section 2.2.1 we have seen that the Fisher index is based on equi-representativity of the basket. We can therefore affirm that the *EKS* index, which gives PPPs that differ only minimally (see 2.1.4) from the PPPs of the Fisher type, is also based on equi-representativity of the basket. Moreover we have seen that the Fisher index is based on a system of common prices equi-distant from national prices and on a set of common quantities equi-distant from national quantities. Note that the indices (0)$G(E)$ and (1)$G(E)$ on one side and the indices (0)$G(O)$ and (1)$G(O)$ on the other side use a set of common quantities obtained as an unweighted average of prestandardized quantities and a set of common prices obtained as an unweighted average of national prices. The explicit use of the unusual process of prestandardization of the quantities in the definition of common quantities and the exclusion of quantity weights in the definition of common prices are explained purely by the resolve to achieve the two-fold objective of abandoning regional typicality and atypicality, which would otherwise be inevitable, and of attaining equi-representativity of the composite item.

The first objective is obviously achieved, which in itself is enough, as we shall see, to make indices (0)$G(E)$ and (1)$G(E)$ different in terms of properties satisfied from index (1)$G(EW)$, and indices (0)$G(O)$ and (1)$G(O)$ different from the *GK* index. The second objective, however, can be claimed to have been perfectly achieved only when no more than two countries are considered. In the general case of more than two countries, the achievement of equi-respresentativity of the composite item is thwarted by what could be called the effect of "similarity of countries." For the indices mentioned above of the E and O type respectively, this means that the higher the number of countries whose consumption patterns (price system) is relatively similar to that of a given country, the relatively more (less) representative the composite item will be of that country. The distance between the multilateral and binary results obtained with the two kinds of indices gives an idea of the "similarity of countries" effect. However, as we shall see later, this difference can be reduced without jeopardizing the fundamental characteristics of these indices, if the prestandardized national

quantities and the national prices are multiplied by special country weights compiled for the precise purpose of achieving equi-representativity.

2.2.5. Impact of the Different Types of Representativity on the Results of the Comparison

In the preceding paragraphs we have seen that the results of the comparison depend on the relative representativity of the composite item. If an index is based on equi-representativity of the composite item, when this index is applied at different geographical levels the relative representativity for a given pair of countries is constant and therefore the results obtained in a bilateral, subregional (e.g. EEC), regional (e.g. Europe), or world context are almost the same (the differences can be seen as random errors).

A comparison based on equi-representativity of the composite item guarantees the stability of the results. The regional typicality and atypicality are obviously conditioned by the characteristics of the region as a whole. This means that, for a given pair of countries, the relative representativity and therefore the results obtained at the different geographical levels differ because the characteristics of the region in which this pair is inserted obviously differ. Moreover, it should not be forgotten that the results obtained applying in successive years indices based on regional typicality or atypicality of the composite item (and therefore subordinated to the characteristics of the region) may show, for a given pair of countries, a relative evolution of prices and volumes substantially different from that obtained from the ratio of the two national indicators simply because this pair is inserted in a region for which the number of countries included changes over time. Those to whom these particular aspects of comparisons are important (stability in results obtained at the various geographical levels, time/space consistency) should avoid indices based on the typical item (e.g. $(1)G(EW)$) and the atypical item (e.g. the GK index) like the plague.

3. Conditions Imposed on Comparisons

The choice of the indices to be used in international comparisons of volumes and purchasing powers is governed by a set of requirements imposed on the results of the comparisons, because other more sophisticated procedures are impracticable. An attempt will be made in the sections that follow to describe and comment briefly on the requirements most frequently applied—whether rightly or wrongly—when comparisons are made.

3.1. *Transitivity (or Invariance with Respect to the "Reference Country")*

This requirement is met if, calling the price or volume index I and any three countries participating in the comparison j, h and r, the following equation applies:

$$ {}_rI_j / {}_rI_h = {}_hI_j $$

If this requirement is met, all the information contained in the matrix of comparisons between all the pairs of countries is also contained in any row or column

of the matrix. All the results of comparisons among K countries may thus be expressed in a concise form as $K-1$ comparisons between each of the countries and any reference country. The fact that the condition of transitivity has been fulfilled is, precisely, a guarantee that the results are invariant with respect to the choice of such a country. As for the composite item approach, imposing the condition of transitivity simply means that the comparisons of all the various pairs of countries are based on the same composite item. In contrast to what is often said, the fact that this condition is imposed on a comparison gives no indication whatsoever of the type of index to be used, as every index—if applied in a particular way—can produce transitive results.

3.2. *Invariance with Respect to the "Base Country" (Neutrality of Countries)*

I should like to begin by saying that the significance of this requirement, such as it can be assessed from descriptions in the specialist literature and from the way it is applied in practice, has never been clear to me. Indeed, if the "base country"—according to the current interpretation—is that—as opposed to the reference country—which "stamps its hallmark on all the results of the comparison," it is not obvious how invariant results can be obtained when the base country changes (if we exclude the hypothesis that all countries have the same hallmark). We must, therefore, suppose that when this requirement is imposed (within the framework of the ICP, for example), the aim is to achieve results that are not "stamped with the hallmark" of any of the countries participating in the comparison, or in other words, a kind of "neutrality of countries." Such an aim does seem—despite what was said above—not only practicable but also of value, particularly for the type of comparisons in which we are interested. Nevertheless, I am not altogether in agreement with the way in which the indices and the systems of comparison are divided into those that meet this "neutrality of countries" requirement and those that do not.

To give an example, I believe that a "star" system of comparison using Fisher's ideal index, such as the one carried out on COMECON countries with the Soviet Union as the center, meets the requirement of neutrality far more successfully than when the GK index (as suggested by the ICP) or the $(1)G(EW)$ index is applied to this same group of countries. It seems to me that the first type of comparison produces results that are less stamped with the hallmark of the Soviet Union than the other two types.

I would suggest that equi-representativity of the composite item should be the criterion when choosing a method to meet the requirement of neutrality of countries, even if this does imply the adoption of star systems of comparison (but only those based on indices of the Fisher type). Up to now, these systems of comparison have been opposed by specialists in the field of international comparisons but in my opinion, and as I shall show in this paper, they deserve to be reassessed, at least at the level of basic headings.

3.3. *Average Tests of Volume Ratios and of PPPs*

This requirement is met whenever the volume ratio (PPP) of an aggregate is between the smallest and largest of the volume ratios (PPPs) of the components of the aggregate.

3.4. *Additivity*

This condition is respected if the real value (volume) of an aggregate of a country calculated directly (by converting the nominal value of the aggregate into a unit of reference by applying the specific PPP) is equal to that obtained from the sum of the real values of the components at any level of aggregation, each obtained by applying the specific PPP calculated independently. This condition is indispensible only when the absolute real values are to be published. If publication is limited to the volume ratios of each country relative to a reference country (the USA in the ICP framework), or relative to a group of countries (EUR 9 in the EEC framework), it is not indispensible. It is nevertheless important, because it is a sufficient condition for ensuring that the volume ratios meet the average test. The only index of those mentioned in section 3 that is additive is the $(0)G(O)$ index.

3.5. *Sectoral Independence of the PPPs and Volume Ratios*

The PPPs and volume ratios relating to a given aggregate must be a function of prices and quantities (volumes) observed in the participating countries for the items of the aggregate. All the indices meet this condition when applied correctly, since it is associated with the procedures of application rather than the indices themselves.

In the ICP framework the *GK* index was applied in such a way as to obtain special volumes that would meet the additivity and average conditions, but which are different from the volume ratios obtained by applying the *GK* index independently to each component of GDP.[14] The ICP version of the *GK* method therefore does not fulfill the condition of sectoral independence of the PPPs and volume ratios. The special PPPs and volume ratios relating to a given aggregate are also a function of the prices and quantities of the items of other aggregates. This means that the PPPs and volume ratios of a component of GDP (e.g., investment goods) cannot be calculated until the prices and quantities for all the items comprising GDP (including, e.g., food, clothing, etc.) are available for all countries. Moreover, possible errors in the prices or quantities for items belonging to certain components, but produced only in some countries, may have significant repercussions on the PPPs and volume ratios of all the components, in such a way that their correction causes changes throughout the whole range of GDP.

3.6. *Proportionality of Quantities*

This condition is respected if, when the quantities of a country are multiplied by a factor, the new volume ratios between that country and the other countries are equal to the old volume ratios multiplied by the same factor. If this condition is met, it then becomes possible to calculate the per capita volume ratios directly by using the *per capita* nominal values, or indirectly by dividing the global volume ratios by the respective populations. If, for example, during a periodic revision

[14]These differences can sometimes be very important. In the first phase of ICP, a difference of 8 percent was obtained in the comparison of investment between the Federal Republic of Germany and the U.S.A., and a difference of 9 percent in the comparison of public consumption between Kenya and the U.S.A.

of the national accounts the GDP of a given country in a given year was increased by 10 percent while retaining an identical structure, it would be expected that if the condition of proportionality of quantities is met the PPPs between the given country and the other countries would remain unchanged, the volume ratios would increase by 10 percent, and the volume ratios among the other countries would remain unchanged.

3.7. *Proportionality of Prices*

A method meets this condition if, when the prices of a country are multiplied by a factor, the new PPPs between that country and the other countries are equal to the old PPPs multiplied by the same factor. If a method meets this condition, the volume ratios are not influenced by expressing the prices of a country in its monetary unit, fractions of that unit, or multiples of it.

3.8. *Applicability in the Case of Negative Balances*

Some components of GDP (the export–import balance, the change in stocks) can be negative, and in practice frequently are. In this case some indices (e.g. *GK*) are inapplicable without resorting to expedients of the type adopted in ICP.

4. Choice of Methods

4.1. *Basic Headings*

At the basic headings level, the indices of choice for comparing volumes and for comparing purchasing powers are the same. The requirements that are relevant at this level are transitivity, proportionality of prices and quantities, and neutrality of countries. The advantages of imposing these requirements have already been pointed out. The remaining conditions discussed in the previous section are aimed at ensuring that the results obtained at a given level of aggregation are consistent with those obtained at lower levels of aggregation, and therefore are not applicable at this lowest level. As noted above, the condition of transitivity does not discriminate among the indices examined. The requirement of proportionality of prices is also met by all of the indices. However, the requirement of proportionality of quantities will entail the elimination of the homogeneous weighted indices (*OW*) and the heterogeneous weighted indices (*EW*). To ensure neutrality of countries, the index selected will have to be based on equi-representativity of the composite item. The indices which provide the best guarantee of equi-representativity are the balanced indices (*B*). The choice of an index within the balanced indices group will depend on the nature of the raw material (prices and weights) actually available.

The first point that needs to be made is that in order to apply a balanced index of the *G* type (e.g., (0)*G*(*B*)), a common basket of commodities is required, for which prices and weights will be needed for all countries taking part in the comparison. In fact such information is rarely available even for relatively homogeneous countries such as those in the EEC. Accordingly, a balanced index of the *VY* type is preferable because such indices are based on matrices of

binary PPPs that can be calculated on the basis of binary baskets. Furthermore, it may be observed that for certain basic headings and specific pairs of countries characterized by different patterns of consumption, there is no single item in the basket for which prices are available in both countries, unless the criteria of equivalence (definitions) of commodities are widened to the extent that systematic differences of quality will distort the comparison. This means that the matrix of binary PPPs that can be calculated is incomplete. In summary, the only balanced index that can be applied without drawbacks when the raw material is of this type is the *EKS* index.

4.1.1. Organization of the System of Comparisons

The fact that the *EKS* is applicable to incomplete PPP matrices of the Fisher type (in such cases it often needs to be repeatedly applied until the matrix has been completed) is a matter for reflection. Indeed, if we accept the idea (widely adopted in temporal comparisons) that a direct comparison of two countries (periods) which are very different is not necessarily better than an indirect comparison carried out using an intermediate bridge country (period), we can understand that there are some types of comparison—apart from direct ones that are impossible to perform—that ought to be discarded. The estimates would then be more precise and the costs of the comparison would be reduced, as we shall see more convincingly when we go on to talk about the strategy of choosing commodities. But those people who are convinced that direct comparisons imply a high degree of precision whereas indirect comparisons imply rough-and-ready measurements, and who therefore favour direct comparisons whenever these can possibly be carried out, are doomed to disappointment, since a calculation of the *EKS* type is a function of all possible indirect comparisons.

Consequently, I believe that the application of the *EKS* to the most complete possible PPP matrix of the Fisher type (the so-called "maximum scale" adopted by the EEC for the 1980 comparison) must in future be replaced by its application to the most incomplete possible PPP matrix of the Fisher type ("minimum scale" procedure). I will not beat about the bush but instead go straight to the point and say that our objective must be to identify the $K-1$ most suitable direct parities, on the basis of which the most precise and complete matrix of transitive PPPs will be obtained, if necessary by applying the *EKS* repeatedly. Depending on the group of countries concerned and on the basic headings, the $K-1$ direct comparisons between the K participating countries will take the form of either a chain system or a single-centre star system (as in the comparison between socialist countries and Austria under Phase IV of the ICP), or a polycentric star system (a highly effective system for world comparisons).

4.1.2. The Regionalization of Comparisons

Once the PPP matrices at the regional level have been completed for each of the basic headings of the world classification, they can constitute the sub-matrices of the grand matrix on a world scale, this work being coordinated by regional bodies. All that then needs to be done to complete the world matrix for each basic heading is to carry out the minimum number of direct binary interregional comparisons considered most suitable on the basis of the baskets

of binary commodities. This small amount of work is all that needs to be done by the coordinating body at world level. A star system could be adopted to carry out direct interregional comparisons. For example, the U.S.A. could be one centre, and the various points of the star could be the Federal Republic of Germany and/or Austria for Europe, Colombia for Latin America, Japan for Asia, Kenya for Africa, and so on.

In the process of obtaining the complete and transitive matrix at world level by means of the (if necessary repeated) applications of the *EKS*, it is possible that the intraregional results will differ slightly from those obtained at the regional level. These differences must, however, be attributed to random errors that tend to balance each other out right from the first levels of aggregation. The use of the special techniques that have been devised to ensure that intraregional results for the basic headings of the world classification remain constant at the world level, even in the presence of several interregional comparisons, does not give any appreciable quantitative advantage at the aggregate level. I do not therefore recommend them.

This way of proceeding, as explained above, is the most accurate one. But even if for the sake of argument this were not the case, it would still be the only method which, given its simplicity, would enable us to obtain the results of a world comparison within reasonable time limits. It is in the interest of sub-regional and regional bodies and of individual countries to advocate a regionalization of world comparisons such as has been broadly described above, if they wish to avoid—and the risk exists—sub-regional and regional comparisons being made again from scratch when a world comparison is carried out on the basis of broad definitions designed to make direct comparisons of the largest number of countries in the world. The existence of other results that—on account of the broad definitions used—are certainly of inferior quality to, and probably quite different from those already existing, is a serious drawback for intraregional comparisons. Some problems also exist for interregional comparisons. Indeed, if the definitions are broad, the fact that each country uses its own local standards—even while remaining within the definition—may give rise to distortions that jeopardize the validity of the comparison between countries in different regions.

Once the transitive PPP matrices for basic headings in the world classification have been completed in the way indicated above, it only remains to apply the chosen aggregation method. The results relating to aggregates obtained at world level will not differ greatly from those obtained at regional or sub-regional level if a suitable aggregation method is chosen, as we shall see in the following section. Just as with the basic headings, so at the level of aggregates differences in the results obtained at the various geographical levels (binary, sub-regional, regional and world) may be regarded as random errors in the measuring process, provided that the index applied at the various levels is of the type indicated below. (For a given pair of countries, the result obtained at world level may be closer to the binary level than the sub-regional level, for example.)

The fact that the same method is used, without regard to its nature, at the various geographical levels, far from being as has been claimed a source of consistency in the results may on the contrary be a source of substantial

differences. This is particularly likely to happen if the method is based on the characteristics (prices or quantities) of the region, which is precisely the case of the *GK* and (1)*G*(*EW*) indices, as these characteristics are understandably subject to significant changes between one geographical level and another. By way of illustration, an India/Luxembourg volume index of the *GK* type that is computed in a binary context (which is more or less the same as the Paasche index) is probably much closer to the volume index computed at the world level with the *EKS* index than that computed with the *GK* index. As world prices are closer to those in rich countries—and hence to those in Luxembourg—the world volume index of the *GK* type is further from the Paasche index—*GK* binary— than from Laspeyres and hence from the *EKS* index, which is nearer the middle of the Laspeyres–Paasche spread (Fisher). We may observe by way of conclusion that the techniques devised to ensure that the results obtained at sub-regional level are maintained constant at regional and world level, and which are sometimes proposed as solutions, should be totally discarded in our opinion, especially at the aggregate level. Fortunately, this opinion seems to be shared by almost all experts in international comparisons. If care is taken to choose the correct method of aggregation, such disparities, which are to be considered as random errors of measurement, are negligible and in any case much less significant than those occurring in the estimates of nominal values; what is more, they can easily be explained to anyone who is particularly worried by such discrepancies. Secondly, special techniques devised to guarantee the fixity of results have their own disadvantages that have been well described in the specialist literature, which jeopardize the validity of world comparisons.[15]

4.1.3. The Selection of Commodities for the Basket

In order to shed more light on the arguments put forward under 4.1.1., it is useful to consider how commodities are chosen for the basket used for calculating PPPs of the Fisher type and therefore of the *EKS* type. It is well known that the stricter the criterion of equivalence of a commodity, the more its degree of representativeness will differ in the various countries. It must be borne in mind that even when weighting factors are applied to each item, the basket of commodities cannot be reduced to an equi-representative composite item if all the items in the basket are more representative of one of the countries than the other. Using price ratios for five French cheeses, for example, it is very difficult to obtain an accurate estimate of Fisher's index for the basic heading "cheeses" when comparing France and the United Kingdom, even if individual weights are applied. This means that it is only possible to obtain an accurate estimate of Fisher's index if there is at least one item in the common basket that is sufficiently representative of consumption in the reference country, and at least one other that is sufficiently representative of consumption in the partner country. Each country ought therefore to put forward one commodity (or several) which it judges to be adequately representative of its consumption to be used

[15]See: L. Drechsler, "International Comparisons of Economic Development Levels—General Review and Selected Problems", Economic Commission for Europe, Geneva, May 1979; I. Kravis, A. Heston, and R. Summers, "The Role of Regionalisation . . .", ECIEL, Rio de Janeiro, January 1979.

in the calculation of the Laspeyres Index based on that country. This strategy for selecting commodities was adopted by the EEC for the 1980 comparison (unilateral approach). The sum total of such commodities constitutes the total basket for a given basic heading. When this has been established, the next step is to decide on a system of comparison.

(*a*) *Maximum scale.* If all the countries taking part decide (and are able) to record prices and obtain weights for all the commodities, it then becomes possible to calculate the complete matrix of Fisher indices. If, as is more probable, weights per commodity are not available with any degree of precision, it is preferable to abandon them, but not overtly to favour non-weighting (equi-weighting) as this would not guarantee the equi-representativity of the composite item. This is valid whether the *EKS* is applied in a traditional manner to Fisher indices or to a non-weighted (equi-weighted) geometric mean of price ratios for common commodities. In order for the composite item to be equi-representative without using commodity weights, the estimate of the Laspeyres index for country j on the basis of country h to be used for the *EKS* index must be obtained by dividing the price recorded in j for the proposed commodity from country h (P_{hj}) by the price recorded in h for the same commodity:

$$_h\hat{L}_J = P_{hj}/P_{hh}$$

This comes to the same thing as the convention of attributing all the weight to the commodity (or commodities) proposed by country h when calculating the Laspeyres index. In such a case, the *EKS* procedure produces a PPP between countries j and h that may be expressed as follows (provided that each country has proposed only one commodity:

$$_h\hat{EKS}_j = \left\{ \prod_{\alpha=1}^{k} \left[\left(\frac{P_{\alpha j}}{P_{\alpha h}} \right) \cdot \left(\frac{P_{h\alpha}}{P_{hh}} \cdot \frac{P_{jj}}{P_{j\alpha}} \right) \right]^{1/2} \right\}^{1/k} = [(0)\widehat{VY}(E) \cdot (0)\widehat{VY}(O)]^{1/2}$$

From this formula it can be seen that the option of equi-weighting associated with the use of the geometric mean and represented—given the strategy used to select commodities—by $(0)VY(E)$, fails to meet the objective of equi-representativity of the composite item insofar as—and this is important at the basic headings level—the "similarity of countries" effect plays a role (see section 2).

Minimum scale. If the more suitable choice of $K-1$ direct comparisons of the Fisher type provides a star system of comparison with a single centre, the country that serves as the centre for a given basic heading will have to supply prices for all the commodities in the basket, while the other countries, in addition to providing the prices of the commodities they themselves have suggested, will also have to provide the price of the commodities suggested by the centre country. In a case such as this, the cost of the comparison to the centre country is identical to that it would incur if a maximum scale were adopted, while for the other countries the cost is appreciably lower. The choice of the most suitable direct comparisons may therefore result in a pattern of comparisons that differs from the star system, and at the same time results in costs being divided more equally among the countries taking part in the comparison.

4.2. *Aggregates*

At levels of aggregation higher than basic headings, the indices suitable for comparing volumes and those suitable for comparing purchasing powers are different, given that the requirements to be met by the results of the comparison are not the same. The requirements of each type of comparison are described below, as are the indices that meet these requirements most satisfactorily.

4.2.1. Comparison of Volumes

The requirements for this type of comparison are transitivity, neutrality of countries, additivity, average test of quantities, sectoral independence of PPPs and of volume ratios, proportionality of prices and quantities, and applicability in the event of negative balances. The imposition of the average test of quantities, a vital requirement for levels of aggregation higher than basic headings, entails the abandonment of balanced indices and therefore of the *EKS*.[16] Of the indices previously examined, the only ones that meet the average test requirement are those which permit a comparison of volumes by means of a "constant price" procedure (in the spatial sense). Of these, the *GK* and those of the $(X)G(OW)$ type must be excluded as they are based on the concept of a composite item that is atypical for the region and not on that of an equi-representative composite item as is necessary to meet the requirement of the neutrality of countries. In addition, these indices must be rejected because they do not meet the requirement of proportionality of quantities. The only choice remaining is among indices of the $(X)G(O)$ type. Our choice among these is further governed by the requirements of additivity and sectoral independence of PPPs and volume ratios. We are left with the $(0)G(O)$ index, which was used in the 1975 EEC comparison and which is known by my name in the specialist literature. This index meets not only the requirements of the proportionality of prices and quantities but also that of applicability in the event of negative balances. As noted above, the $(0)G(O)$ index, like the other indices of the $(X)G(O)$ and $(X)G(E)$ type, completely guarantees the neutrality of countries only when only two countries are being compared. In more general cases, the attainment of neutrality of countries may be hampered by the similarity of countries effect.

In my experience, the negative impact of the similarity of countries effect is fairly limited. It could be reduced somewhat by carrying out a preliminary stratification of countries on the basis of a preselected criterion of similarity. By applying an equi-weighting between the strata and than an equi-weighting within the strata, special country weights may be derived to be attributed to prices in the various countries when defining the common structure of prices associated with the $(0)G(O)$ index. If one is prepared to forego the sectoral independence of PPPs and volume ratios, other solutions are feasible. Special country weights of the type mentioned above, for instance, can be computed by means of numerical procedures in such a way as to minimize the differences—defined

[16]In an extreme case where quantity ratios between two countries were the same for each commodity, the Fisher volume index would be equal to this constant at each level of aggregation. The same cannot be said, however, for the *EKS* index, which is outside the interval (nil in this case) between the smallest and largest of the basic volume ratios. When countries are similar to each other, it is more likely that the average test of quantities will not be met by the *EKS* index.

according to a preselected criterion—between transitive estimates and estimates obtained, for a preselected level of aggregation (such as the overall level) by bilaterally applying the $(0)G(O)$ index. A possible option, if sectoral independence is not required, would seem to be to choose the value of $-\infty < X < +\infty$ for which the absolute minimum is obtained. However, as the value of X varies from one comparison to another, this solution is of little practicable merit, and it is better to choose $\lim X \to 0$, i.e. the geometric mean. Apart from the definitions of common prices that are formed from the weighted averages of national prices (which themselves constitute a subgroup of all possible definitions), it is possible to devise other definitions of common prices that are equi-distant from national prices.

To sum up, whatever the type of definition adopted, it is important to accept the idea that common prices must be equi-distant from national prices. The notion of common prices that are equi-distant from national prices, which I am putting forward as an alternative to the average prices of the group of countries (GK) used up to now for ICP, will not be accepted until the following points have been made clear:

(i) The common prices used in the "constant-price" procedures to measure volumes in the various countries are like an elastic tape measure, the length of which is in inverse proportion to the "distance" between such prices and prices in individual countries.

(ii) "Equi-distance" of common prices with respect to national prices is the only way to "stiffen" the tape measure and hence the only way to eliminate distortions in volume measurements and guarantee the stability of the results obtained at the various geographical levels (binary, regional and world), in the sense that it reduces the remaining differences to the level of mere random errors.

4.2.2. Comparison of Purchasing Power

The requirements to be met by this type of comparison at aggregation levels higher than basic headings are transitivity, neutrality of countries, the average test of PPPs, sectoral independence of estimates, and proportionality of prices and quantities. The imposition of the average test of PPPs, which is a vital requirement in this type of comparison, obliges us to abandon balanced indices and hence the *EKS*, and to choose an index from those examined earlier that permits price levels to be compared by means of a "constant quantity" procedure. Among these, those of the $(X)G(EW)$ type may be excluded as they do not meet the proportionality of quantities requirement and are based on the concept of a typical composite item for each region and not on an "equi-representative" item such as is required if the neutrality of countries is to be guaranteed. Our only recourse, therefore is to indices of the $(X)G(E)$ type. The need for the estimates to be sectorally independent compels us to choose the $(0)G(E)$ index from among these. This means that the common quantities used to measure relative price levels have to be computed as a geometric mean of national quantities. As already noted, the $(0)G(E)$ index satisfactorily meets the requirement of the neutrality of countries when only two countries are involved, but in multilateral operations the attainment of this objective may be hampered by

the similarity of countries effect. However, it is possible to obtain common quantities that are equi-distant from national quantities in the same way as it is possible to obtain common prices that are equi-distant from national prices in the ways described earlier.

COMMENTS ON: "SELECTED PROBLEMS OF INTERCOUNTRY COMPARISONS ON THE BASIS OF THE EXPERIENCE OF THE EEC"

BY IRVING B. KRAVIS, ROBERT SUMMERS, AND ALAN HESTON

University of Pennsylvania

The paper by Gerardi covers considerable ground, touching on a wide variety of issues in the area of international comparisons of product and purchasing power. Since our views on most of these subjects have been expounded in one or another of the International Comparison Project volumes and papers[1] we will concentrate mainly on the central issue raised by Gerardi of the selection of an aggregation process that must somehow take account of the tastes of all the people who are the subject of an international comparison inquiry. In addition, we comment on some other points including the notion of special purpose PPPs. Finally, we make a brief statement about where we think future research will be most useful in improving international comparisons.

AGGREGATION IN INTERNATIONAL COMPARISONS

Unless international comparisons are generated from a model founded explicitly in the pure theory of economic behavior,[2] some way, explicit or implicit, of assigning relative valuations to different goods must be found. To arrive at such relative valuations, one must decide how to translate diverse relative prices in different countries into a common set of weights by which the quantity bundles of all countries may be judged. (In the pure theory of consumer behavior one would eschew the assignment of *any* mixed set of common prices as a basis for relative evaluation of different goods because marginal utilities are not the same at different consumption points.) Gerardi's translation averages relative prices across countries, treating each country as a unit equal in importance to each other, irrespective of size or level of *per capita* income. The ICP's translation on the other hand attempts to the extent possible, through supercountry weighting and the technical features of the Geary–Khamis method, to treat each transaction (that is, the purchase of a unit of final good) as a unit equal in importance to each other like transaction, irrespective of the size of the country in which it is carried out.[3]

[1]Kravis, Kenessey, Heston, and Summers (1975); Kravis, Heston, and Summers (1978); Kravis, Heston, and Summers (1982).

[2]Most empirical studies, but by no means all, drawing on price and quantity data of different countries have had a demand function orientation rather than an international comparison objective. Chapter 9 of Kravis, Heston, and Summers (1982) contains a generalized demand system study directed explicitly at the problem of international volume comparisons.

[3]In the unit-country weighting method favored by Gerardi the comparison between France and Germany is as much influenced by the price structure confronting the 400,000 inhabitants of Luxembourg as it is by the price structure confronting more than 50 million people in France and more than 65 million in Germany. For the set of 9 countries, "characteristicity" is diminished for many more people than the number for which it is increased.

The method favored by the ICP (and endorsed by Professor T. P. Hill in a methodological report commissioned by the European Community, the Economic Commission for Europe, and the UN Statistical Office[4]) is in accord with standard national accounts practice in each country; the relative prices are the international analogues of the national average transactions prices that are implicitly embedded in each country's national accounts expenditures. The ICP method also produces the essential consequence that the resulting prices and quantity comparisons are invariant to changes in the political subdivisions of the region or world into national states.[5]

If in fact we are trying to capture average world prices, then account must be taken of the number of transactions. If there are more transactions (as there would be if the population of a country doubled and all else remained the same) then we should expect the average prices and the resulting index numbers to change. Thus we think Gerardi's proportionality criterion is inappropriate.

At a number of points, including his arguments for preferring EKS over CPD, Gerardi assumes an inherent superiority in Fisher "ideal" indexes. As Hill points out, "When the group approach [i.e. an approach involving more than two countries] is used no special significance is to be attached to the results obtained from isolated binaries: in particular, the binary results are not to be elevated to the status of norms by which to evaluate or appraise the multilateral results. From a theoretical point of view there is no sense in describing a particular multilateral index as biased because it happens to be close to a Laspeyres index. The attribution of bias in these circumstances is equivalent to asserting that the index should have been defined with respect to a typical consumer from one or another of the two countries in question instead of a representative group of consumers. But this is merely to reassert the difference in value judgments from which the two different approaches stem." (Hill, 1982, p. 31)

Some Other Items

There are a number of other assertions made by Gerardi which in our view represent errors in judgment or fact. Gerardi's advocacy of the average volume test[6] is not necessarily as attractive as he suggests. For example, it can be shown that in the context of the theory of consumer behavior the true cost-of-living quantity index need not lie between the maximum and minimum quantity ratios. In the parallel case of the average PPP test, Gerardi's statement that the ICP's PPPs fail to meet this test ignores a critical feature of the Geary–Khamis system. If a country's summary category PPPs are expressed relative to the world as a

[4]T. P. Hill (1982).

[5]The Gerardi method would assign the same weight to Luxembourg and Belgium prices as to German and Netherlands prices in a comparison involving the four countries. However, if Luxembourg and Belgium became one country their average prices would have a combined weight of one. The comparison between Germany and Netherlands would differ according to whether Luxembourg and Belgium were treated as two countries or one.

[6]The test holds that volume comparison for an aggregate should lie between the minimum and maximum of the components of the aggregate; that is:

$$\min_i \frac{q_{i3}}{q_{i2}} < v < \frac{q_{i3}}{q_{i2}}.$$

whole rather than relative to a numeraire country, the PPP for an aggregate would necessarily lie between the smallest and largest of the PPPs of the components of the aggregate.[7]

It is puzzling why Gerardi objects to an aggregation method that requires special treatment for negative items (e.g. inventory changes). All of the commonly considered methods are designed to compare physical volumes. It is not to be expected that without appropriate adjustments they can be routinely applied to net items in the national accounts that are different in character from the physical flows of the other components of final expenditures on GDP.

Finally, Gerardi contends that GDP purchasing power parities should not be computed simply as a dual emerging from the Geary–Khamis system. We consider the basic PPPs presented in the ICP report (and also in the EEC studies), which refer to GDP as a whole, to be the most useful for general purposes. At the same time the ICP provides a framework for the computation of a comprehensive set of PPPs for any desired subaggregates. The ICP's Phase III report contains supplementary estimates of PPPs for private households and suggests methods of deriving other special purpose PPPs such as, for example, those pertaining to the real earnings of any particular group of workers. These would require relatively modest supplements to ICP data. Apparently Gerardi would elevate the importance of the PPP used to compare real earnings, and for some particular concerns that PPP is surely the relevant one. But for general purposes the PPP covering all of the output of GDP is the appropriate one.

Future Prospects

The methods used in the first three phases of the ICP doubtless will be improved as time goes on. In our view the most promising potentials for substantial progress lie with better sampling of comparative prices and further work on comparisons of services. The pay-off from improved aggregation methods is likely to be much smaller. Plausible alternatives experimented with thus far do not appear to make very much difference in most cases.[8] However, a key criterion for improvement in aggregation methodology would be the

[7]For consistency the summary category PPPs published in the ICP reports are presented with U.S. equal to one, in which case the average of such PPPs need not be bounded by its components. This is because a U.S. summary category PPP *vis a vis* the world as a whole in general is not equal to one. (Only for all of GDP is it equal to one.) Consider the detailed category PPPs of a country, each expressed relative to the world as a whole. The country's PPP for a summary category will be an appropriately weighted average of the PPPs of the individual categories comprising the summary category. Thus the average PPP test will be met. Only when the PPPs are expressed relative to the U.S. is there an apparent violation. The point can be illustrated by referring to the 1975 DM/$ PPPs for total government and the two government components, compensation and commodities (Kravis, Heston, and Summers (1982), Table 6.3, p. 179). The total government PPP is 3.55 while the component PPPs are 3.20 and 2.98 respectively. When these are adjusted for the U.S. PPPs for government, compensation, and commodities—0.91, 0.71, and 1.12 respectively—the German PPPs relative to the world as a whole line up in accordance with the average PPP test: the PPP for government, 3.91, is between the PPP for compensation 4.51, and the PPP for commodities, 3.66. As is to be expected, the average of 4.51 and 2.66 using the real quantities of German compensation and commodities as weights (I$315.2 and I$149.0 from Table 6.5, p. 187) is equal to 3.91, the total government PPP.

[8]Indeed, the ICP's comparison of the Geary–Khamis results with those of unit country weighting (Gerardi's method) show the differences are between 3 and 12 percent.

establishment of a closer link to economic theory than the still largely empirical Geary–Khamis formulation affords. Gerardi's proposals move further away from economic considerations and are permeated by rather mechanical statistical constructs.

References

Gerardi, D. (1982), "Selected Problems of Intercountry Comparisons on the Basis of the Experience of the EEC", *Review of Income and Wealth*, December.

Hill, T. P. (1982), "Multilateral Measurements of Purchasing Power and Real GDP", *Eurostat*, Luxembourg: Statistical Office of the European Communities.

Kravis, I. B., Z. Kenessey, A. Heston, and R. Summers (1975), *A System of International Comparisons of Gross Product and Purchasing Power*, The Johns Hopkins Press, Baltimore, Maryland.

Kravis, I. B., A. Heston, and R. Summers (1978), *International Comparisons of Real Product and Purchasing Power*, The Johns Hopkins University Press, Baltimore, Maryland.

———, (1982), *World Product and Income*, The Johns Hopkins University Press, Baltimore, Maryland.